AUTOMATIC VOLUME COMPENSATION FOR NOISY ENVIRONMENTS

by

Charles Randall McGough

United States
Naval Postgraduate School

THESIS

Automatic Volume Compensation
For Noisy Environments

by

Charles Randall McGough

June 1969

Automatic Volume Compensation

For Noisy Environments

by

Charles Randall McGough
Lieutenant (junior grade), United States Navy
B. S., United States Naval Academy, 1968

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
June 1969

ABSTRACT

The importance of environmental noises which interfere with the audio output from the loudspeaker of an intercom or radio communications receiver has led to the study of sound masking and noise interference. The results are applied in the development of an automatic control system which compensates for environmental noises by maintaining the speaker output at a prescribed listening level above such interference. Such a system could be used to maintain the readability of communications in spite of a noisy environment such as on the bridge or in the engine room of a ship.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

I. INTRODUCTION

Various methods have been developed to adjust the gain or output of an amplifier with reference to chosen signal parameters within the electronic unit. three major amplifier automatic-gain-control techniques have become commonplace in modern circuit design.

Automatic gain control is used in receiver radio-frequency amplification stages to compensate for varying received carrier strength levels. In such systems a control voltage is obtained which is proportional to the amplified carrier strength in one of the amplifier stages, and this voltage is used to control the gain of one or more of the preceding amplification stages. Intermediate-frequency amplifier stages may also incorporate this same automatic gain control. Using this control technique, the output signal strength from the r-f or the i-f amplifier stages may be maintained at the desired level in spite of input carrier level fluctuations.

Automatic gain control is used in audio-frequency amplification stages to compensate for varying audio input levels. The audio output level of the amplifier is monitored and any variance from the prescribed level, caused by input fluctuations, generates a control signal which corrects for the output variance, thus maintaining a prescribed audio output level. This control method is similar to the carrier control technique, the difference being that the audio output average or average peak signal level is the control parameter instead of the signal carrier level.

Automatic tone control is sometimes used in audio amplifiers to compensate for the varying frequency response of the human ear. This frequency response is a function of the sound power level. The automatic tone control boosts frequencies which are less easily heard and attenuates those frequencies to which the ear is most responsive for the various listening levels, thus maintaining a prescribed ability to hear designated frequencies at various listening levels.

These control circuits are designed to minimize variances due to the changes in the strength of the input signal so that the desired listener response will be maintained with a minimum of manual adjustment. Of the parameters considered, however, none of these systems considers or compensates for environmental noise, and quite often the environmental noise level is the critical factor which determines whether the listener will be able to hear and understand the message from the loudspeaker.

It is the intent of this paper to investigate the parameters which must be considered and automatic control circuits which might be employed in compensating for environmental noises. In this way the listener will be assured the ability to hear and understand the loudspeaker output under varying environmental noise conditions.

II. SOUND MASKING

Since the ultimate objective of the audio output of a radio receiver is to provide the ear with a satisfactory input volume level in spite of interfering noise, it is necessary to determine the volume input level which would be required under different environmental noise conditions to allow adequate intelligibility for the human ear.

Hearing ability itself varies from ear to ear as is shown in the variances of volume level different people require to hear either a sound in a quiet environment or to selectively hear a sound in a noisy environment. Most of these variances are physically accounted for by differences in the number of nerve fibers involved in transmitting nerve impulses from the hearing mechanism to the brain and by differences in the frequency of these nerve impulses carrying the sensed information, [1]. These two main physical differences account for a large disparity of hearing ability among people who are all considered to have normal hearing.

Not only does hearing ability vary from person to person even under identical environmental conditions, but it also varies within the individual depending on the nature of the sound. Generally, a pure tone of fixed intensity will sound loudest for frequencies near the middle of the audio range and will sound weaker as the frequency is either increased or decreased. The rate at which a tone sounds weaker or stronger as the frequency is changed from the mid-audio frequencies varies from individual to individual. Also, the ability of various individuals to hear certain bands or groups of frequencies is greatly reduced, while their response to frequencies on both sides of these bands is entirely normal, [2]. Variances occur therefore not only in how different people hear the same sound, but also in how the same person hears various sounds. An average response is obtainable, however, and is satisfactorily accurate in predicting the nearing ability of the normal person, [3].

The primary concern of this paper is the ability of the ear to discriminate between competing sounds or signals. Extensive testing has been performed by both Fletcher and Be'ke'sy on masking effects, or the ability of one sound to mask the presence of another. In testing with pure, single tones, it was initially discovered that low-pitched tones had greater general masking effects than high-pitched tones. In analyzing this effect, Fletcher first measured the hearing threshold of a certain tone, i.e., the level at which a tone is first heard in a quiet environment. He then measured the hearing threshold of the same tone in the presence of a masking tone. The results of some of these measurements are shown in Figure 1, [4].

In Figure 1, the label at the top of each graph indicates the frequency of the masking tone, and the label above each plotted line indicates the level of intensity of the masked tone, measured in decibels above the pressure level of 0.0002 $dyne/cm^2$. The abscissa indicates the frequency of the masked tone, and the ordinate indicates the decibel shift above the hearing threshold that the masked tone must

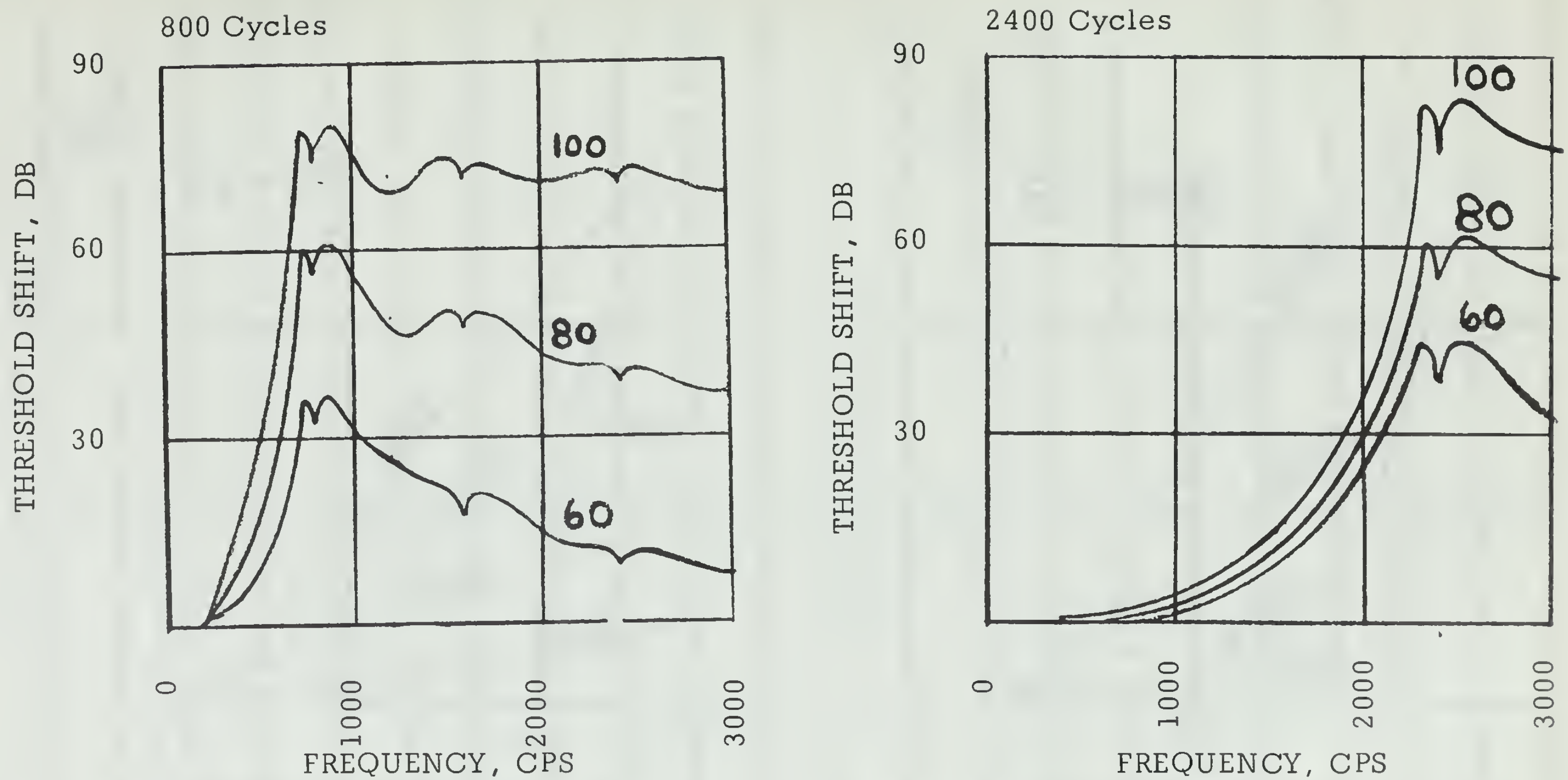


FIG 1. - MASKING DATA FOR 800 AND 2400 CPS. FROM FLETCHER, "SPEECH AND HEARING IN COMMUNICATIONS," D. VAN NOSTRAND COMPANY, INC., NEW YORK, 1958.

be raised in order to become audible again. For example, with a masking tone of 800 cycles per second at an 80-db level, a 400-cycle-per-second tone needs to be shifted 6 db above its hearing threshold level to be heard, but a 1200-cycle-per-second tone needs to be shifted 48 db above its hearing threshold in order to be heard. Notice that frequencies below that of the masking tone may be heard at much lower levels than frequencies above the masking frequency. The results show, then, that a masking tone of a given frequency has a greater masking effect on higher frequencies and has a much smaller effect on appreciably lower frequencies. A tone has its greatest masking effect on the frequencies in the immediate vicinity of the masking frequency.

The same effect is illustrated differently in Figure 2, [5]. The frequency of the masking tone, Fp, is shown at the top of the graph, and its sound level is shown on the abscissa. The frequencies of the various masked tones are shown on each plotted line, and the level the masked test tone must be raised above the hearing threshold level in order to become audible is shown on the ordinate. For example, for a masking tone of 800 cycles per second at an audio level of 70 db, a 250-cycle-per-second tone needs to be only 6 db above its threshold hearing level to be heard, while a 1500-cycle-per-second tone needs to be 40 db above its threshold hearing level. From these graphs it is again seen that a low-frequency tone can completely mask higher

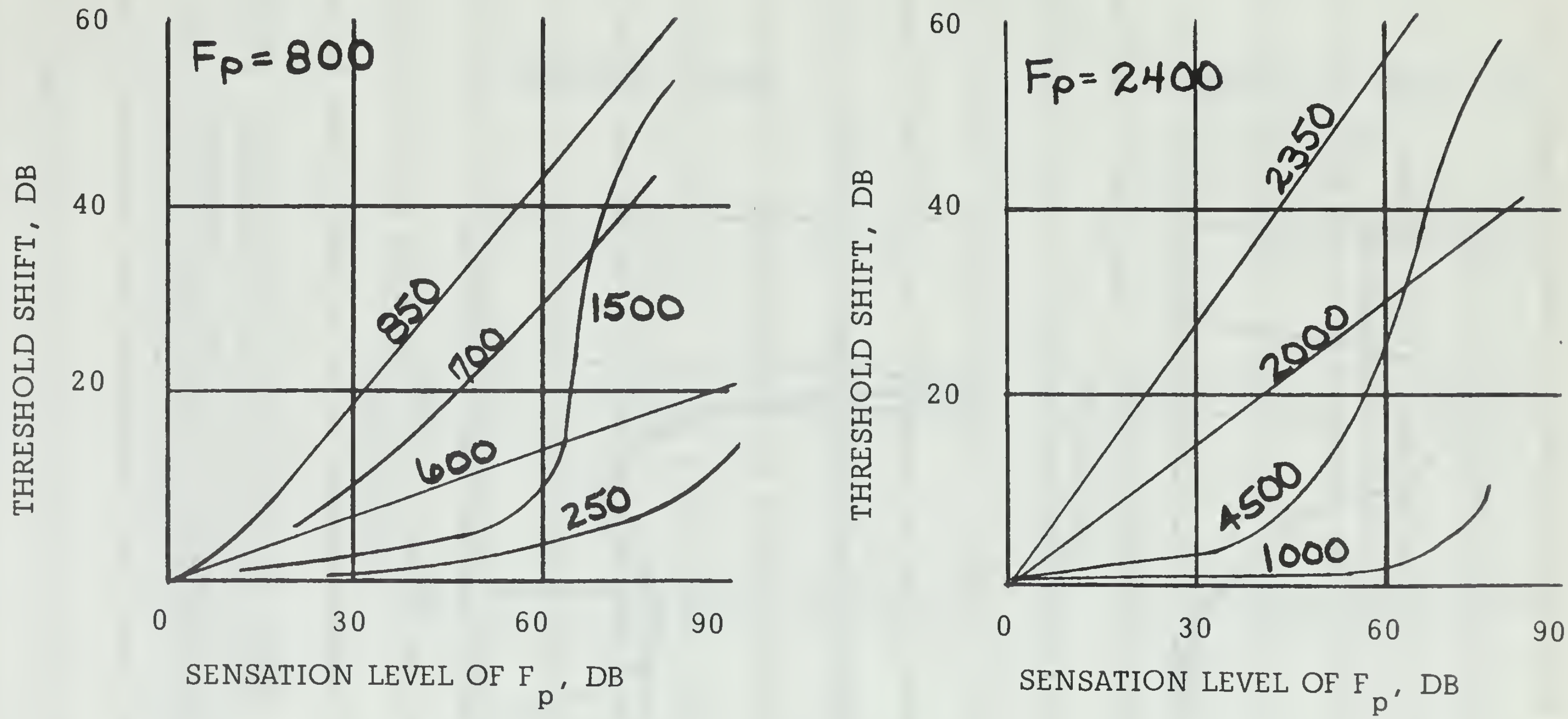


FIG 2. - MASKING DATA FOR 800 AND 2400 CPS. FROM FLETCHER, "SPEECH AND HEARING IN COMMUNICATIONS," D. VAN NOSTRAND COMPANY, INC., NEW YORK, 1958.

frequency tones of considerable intensity, while a higher frequency tone has but a small masking effect on lower frequency tones. A tone masks most effectively near its own frequency.

Experiments were also performed using complex tones, i.e., tones consisting of more than two frequencies. It was found that the complex tone had the same effect as the composite effect of the individual frequency components, [6]. In general, as a complex tone is increased in intensity, the low-frequency tones tend to mask the higher frequencies and are therefore more apparent. Since the ear exhibits a nonlinear response, however, consideration must be given to distortion and to new apparent tones so generated, but this effect is minor and the simple composite effect is satisfactory in predicting response.

Some results from testing the masking effects of complex tones are shown in Figure 3 and Figure 4, [7]. The complex masking tone consisted of tone frequency components 60 cycles per second apart. The graph on the right in each figure shows the sound amplitude pressure of the various frequency components in the complex tone. The graph on the left shows the shift above the threshold hearing level needed for various frequencies to be heard through the complex masking tone. Notice that frequencies lower than those contained within the complex tone are more easily heard because higher frequencies don't mask lower frequencies very effectively. The frequencies higher than the complex tone must be considerably louder in order to be heard, due to the effective masking of the higher frequencies

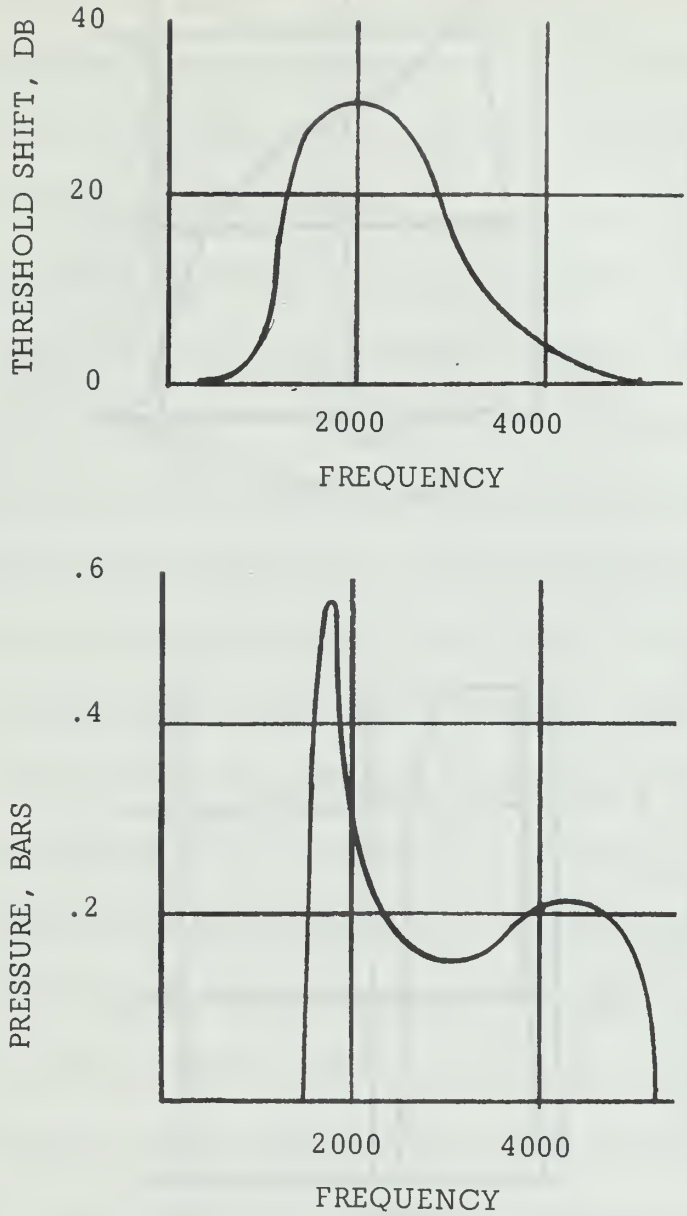


FIG. 3 - MASKING CURVES FOR COMPLEX TONE FROM FLETCHER, "SPEECH AND HEARING IN COMMUNICATIONS," D. VAN NOSTRAND COMPANY, INC., NEW YORK, 1958.

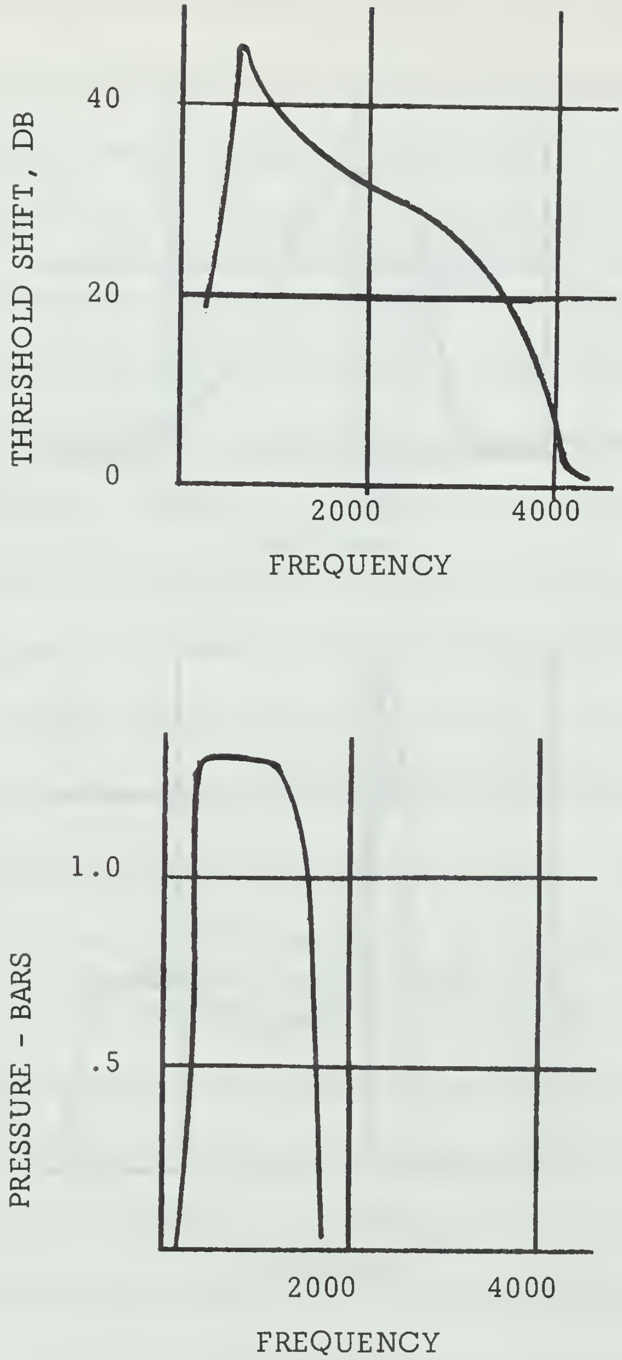


FIG. 4 - MASKING CURVES FOR COMPLEX TONE FROM FLETCHER, "SPEECH AND HEARING IN COMMUNICATIONS," D. VAN NOSTRAND COMPANY, INC., NEW YORK, 1958

by the lower frequencies in the complex tone. Frequencies within the immediate vicinity of the frequencies composing the complex tone are masked most effectively, while those far removed are masked least effectively. It can thus be seen that the masking effects of complex tones and single tones are very similar.

Also tested were the masking effects of band-limited thermal noise, i.e., noise with a continuous frequency spectrum within the band limits. The results are shown in Figure 5 for a noise band between 530 and 1600 cycles per second, [7]. The label on each plotted graph indicates the level of the thermal noise, and the rectangle at the bottom of the graph indicates the band limits of 530 to 1600 cycles per second. The masking effect is very similar to that of both complex and single tones. Although only the results of testing with a thermal noise band of medium width are shown in Figure 5, tests were also conducted with narrow-band and wide-band thermal noise and the results were similar, [8]. The masking effect of the noise was simply the composite masking effect of the frequencies composing the noise. Band-limited noise most effectively masks neighboring and higher frequencies and least effectively masks lower frequencies and frequencies far removed from the frequencies composing the noise.

In all experimentation previously discussed, the test tones were mixed together before being heard by the ear. Fletcher also performed tests in which there was no mixing of the masking or masked tones and where the masking tone was fed into one ear while the test tone was fed into the other ear. Thus complete sound separation was

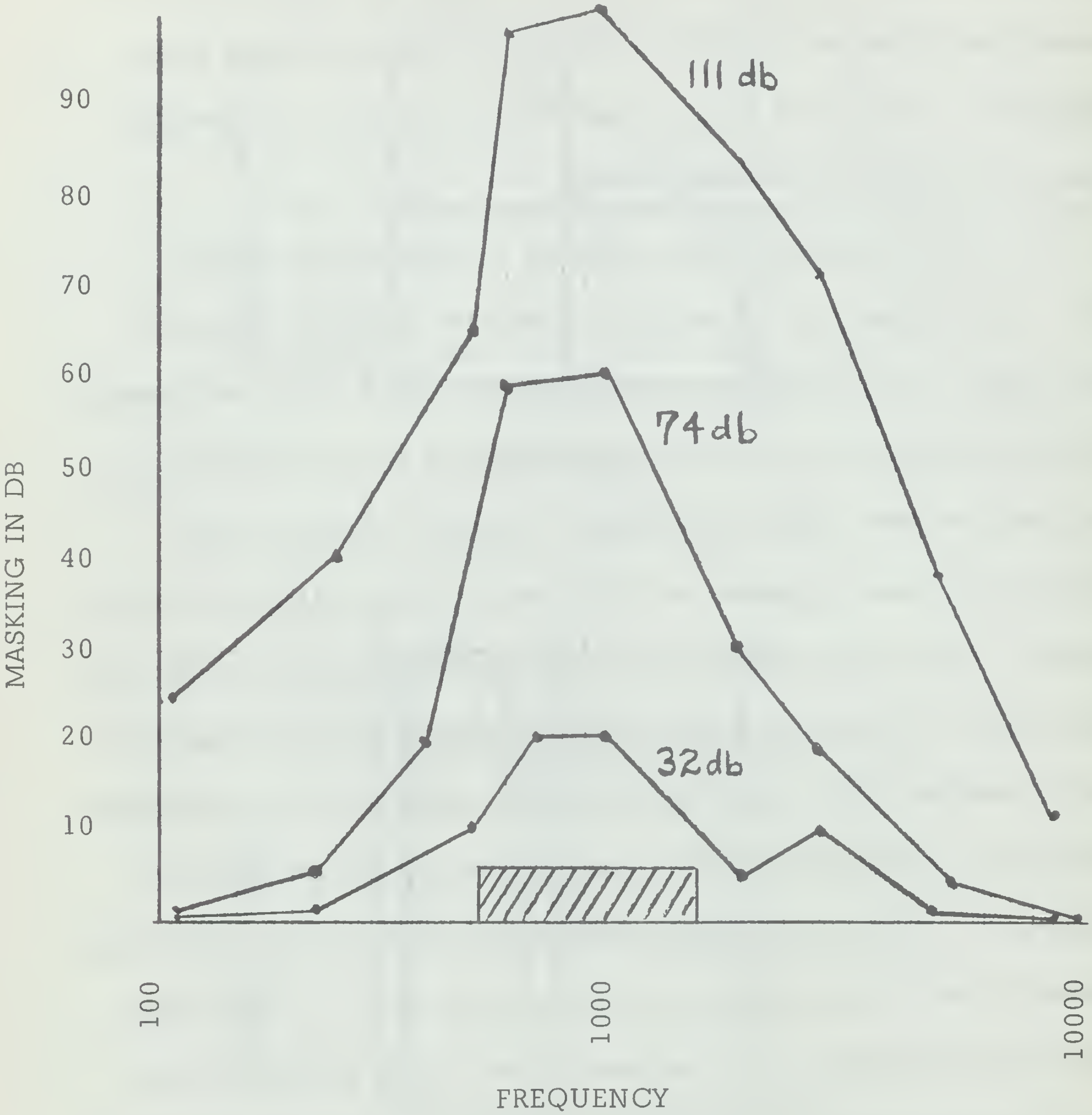


FIG 5. - MASKING CURVES FOR THERMAL NOISE FROM FLETCHER, "SPEECH AND HEARING IN COMMUNICATIONS," D. VAN NOSTRAND COMPANY, INC., NEW YORK, 1958

obtained in what may be called a stereo or binaural situation. It was found that as long as no mixing occurred, masking effects were completely eliminated. Masking and interference were caused only when the competing signals fell on the same ear.

Some of the results of both binaural and monaural testing are shown in Figure 6, [9]. The masking frequency, Fm, and the test frequency, Ft, are shown at the top of the graphs. The abscissa indicates the level of the masking frequency and the ordinate indicates the shift of the test tone level above the audible threshold value necessary for it to be heard. The solid line indicates nonaural testing and the dashed line indicates stereo testing. For example, with a masking frequency of 1200 cycles per second and a test tone of 1000 cycles per second, monaural testing shows that 10 db of masking noise will effectively mask a test tone of the threshold level; but if the masking tone is introduced into the opposite ear, over 70 db of masking noise may be generated before any masking effect is noticed. The only reason why masking occurs at all with binaural testing seems to be because the head bone provides a cross conduction between the two ears, causing some signal mixing at higher audio levels, [10].

Testing was performed by this investigator to verify the preceding results of Fletcher and Be'ke'sy. Instead of using either complex or single tones, however, the masking noise was a variety of commonly encountered noises and the test signal was the numan voice. The two

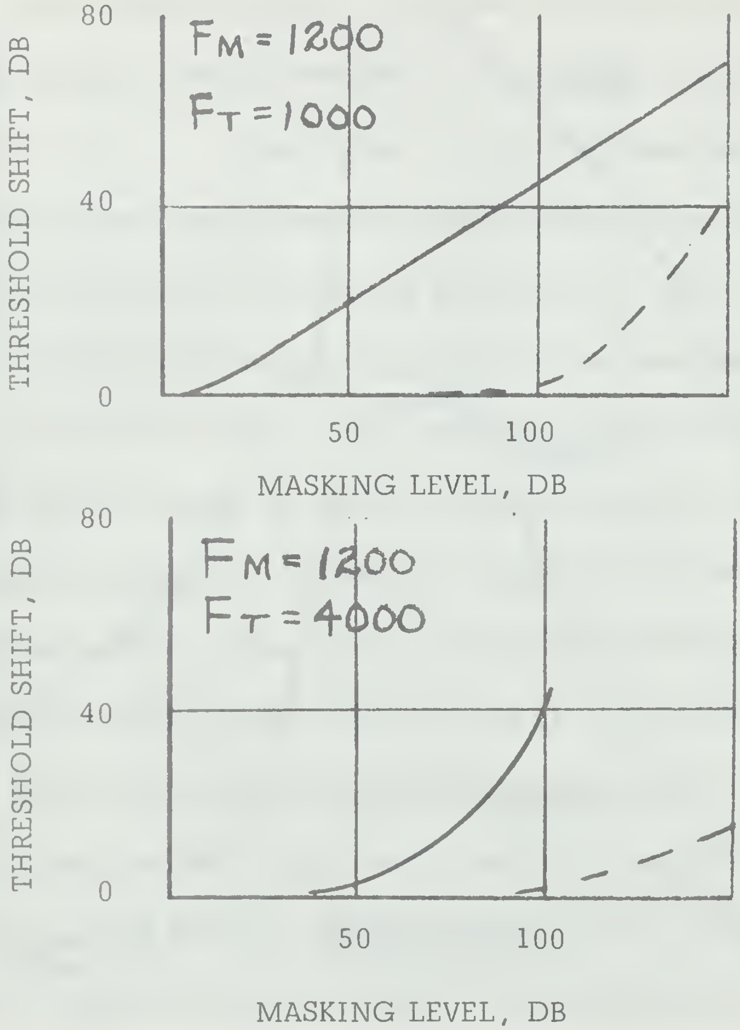


FIG. 6 - STEREO AND NONAURAL MASKING FROM FLETCHER, "SPEECH AND HEARING IN COMMUNICATIONS," D. VAN NOSTRAND COMPANY, INC., NEW YORK, 1958.

signals were fed into loud speakers, and the signal levels recorded were the noise input to the speaker and the voice level at which speech just became intelligible. The results are shown in Table 1.

The data displayed in Table 1 is the average data from the results of testing five persons. The masking noise and the test voice were fed into two separate speakers. The speakers were placed side by side for monaural testing but were moved 45 degrees on each side of the line of sight for stereo testing. The loud speakers in all cases were 10 feet from the listener. The stereo testing was not pure stereo since complete separation of the two sounds was not maintained. The intermixing of the noise and the voice to produce a partial stereo effect is much more realistic of commonly-encountered situations, however, and it is interesting to note the significant difference between the pure monaural and the partial stereo masking effects.

Tests 1 and 2 compare the masking effects of neighboring frequencies to the masking effect of more widely separated frequencies. The noise in test 1, with frequency components close to those of the test voice, masks the test voice much more effectively than the noise in test 2. Much less audio power is required in test 2 to hear the test voice through the noise because the frequency components of the noise are much more widely separated from those of the signal in test 2 than they are in test 1.

TABLE I
EXPERIMENTAL RESULTS OF NOISE MASKING TESTS

NOISE SOURCE TEST	RELATIVE* NOISE LEVEL DB	RELATIVE* VOICE RECOGNITION LEVEL DB	
		MONAURAL	STEREO
1	0	-08	-16
	10	-01	-11
2	0	-12	-22
	10	-07	-18
3	0	-07	-13
	10	0	-08
4	0	-11	-16
	10	-04	-12
5	0	-07	-17
	10	-03	-08
6	0	-07	-14
	10	-01	-06
7	0	-03	-09
	10	+04	-03

*All measurements taken by VOM unmatched to speaker impedance placed across speaker terminals.

NOISE SOURCES:

1. Simulation of rain on sidewalk, noise frequencies close to frequencies in test voice.
2. Simulation of rain on tin roof, noise frequencies much lower than those in test voice.
3. Static from AM radio.
4. Static from FM radio, more constant than AM static, not as pulsating.
5. Background-type contemporary music.
6. Human speech higher in frequency than test voice.
7. Human speech lower in frequency than test voice.

Tests 3 and 4 compare the masking effect of a noise of constant amplitude to that of noise with a varying amplitude, in this case noise with pulsating amplitude. Although the average noise output level was constant, the test voice had to be at a much higher level to be heard through the pulsating noise, test 3, than through the constant noise, test 4. The results from using music as a noise source indicate that music has an effect which varies between that of noise source 3 and 4. With the tremendous variety of modern music, it would most likely be correct to assume that great variation of masking effects would be obtained by using music as a masking noise.

Tests 6 and 7 compare the masking effect of voices of higher fundamental frequency than the test voice to that of voices of lower fundamental frequency than the test voice.

Comparing test 7 to the previous tests, it is interesting to note that the greatest masking effect of the experiment was obtained when the noise and the test signal were similar in content, here both voices. This effect is also evident in test 6, though to a much smaller extent since the noise was higher in frequency than the test voice.

The last two columns in the table compare the results of monaural and stereo testing. In four cases the masking effect was reduced by at least 10 db in stereo testing, and in no case was the effect reduced by less than 5 db. Even though perfect separation was not obtained a significant stereo effect was evidenced, as previously mentioned.

It should also be noted that even with equal noise levels, the degree of sound masking depends to a great extent on the type and content of the sounds being heard. In these tests, the masking ability of a 10-db noise varied from +4 db in test 7 to -22 db in test 2.

These results agree with the basic results of Fletcher and Be'ke'sy. The experimental results discussed in this chapter indicate that even though the masking effect of a noise may depend on many complex factors, basic trends are inherent in noise masking and these trends are predictable.

III. NOISE ANALYSIS AND CIRCUIT REALIZATION

In order to compensate for the masking effect of interfering noises, it is necessary to analyze the signal and the noise in relation to each other. As discussed in the previous chapter, the masking effect of a noise is not only dependent on the level of the noise but also on the frequency content and, more importantly, where the noise frequency components lie in relation to the frequencies contained in the masked signal. It has been shown that noise frequencies lower than or in the vicinity of the signal frequencies have a much greater masking effect than noise frequencies higher than the signal frequencies. Since the frequency relationship effect is greater than the effect of a varying amplitude, consideration must be given to frequency analysis.

The importance of frequency analysis in masking effect may be shown from consideration of Figure 1. It is seen that a 100-db noise tone of 2400 cycles per second causes a threshold shift of 6 db for a 1000-cycle-per-second tone, while a 100-db tone of 800 cycles per second causes a much greater threshold shift of 75 db for the same 1000-cycle-per-second tone. As a result, a much greater compensation is needed for the lower-frequency masking tone in order to maintain the intelligibility of the test signal.

In order to weight the noise frequencies for accurate compensation, it is necessary to analyze the signal of interest, determine the highest frequency components it is desired to detect in the signal, and then to compensate for all frequencies of noise in the vicinity of and lower than those signal components. The exact masking effect of these lower frequencies will differ for varying combinations of frequencies and modulations, as shown in Table 1, but in most cases the signal listened to will be at a level well above the threshold hearing level, and thus any small variation in the low-frequency masking effect will not affect the intelligibility of the signal once the initial compensation is determined.

In weighting the compensation for higher frequencies, it would be more desirable to overcompensate by overweighting frequencies which are less effective in masking ability since the only result would be to have the signal stronger than is necessary when those frequencies of noise are present. Undercompensation for possible masking frequencies would be risking a loss of intelligibility should these noise frequencies appear. Too much overcompensation, however, would result in uncomfortably loud signal levels. It is necessary, therefore, to compensate for noise frequencies slightly higher than those expected to have a masking effect and then filter out those still higher frequencies which can be assured to have little masking effect. This function could be performed by a compensated amplifier sensitive only to frequencies lower than the highest frequency component of noise expected to have a masking effect on the signal.

The two audio reproductions most commonly listened to are the human voice and music. Music may also contain the human voice. These two types of signals should be analyzed separately to determine frequency-masking relationships.

Intelligibility in the average human voice is contained largely in the frequencies from about 500 to 2500 cycles per second, [11]. The results of the experiments reported in the following chapters indicate that noise-frequency compensation should taper off considerably for noise frequencies greater than about 5,000 cycles per second. This allows adequate compensation for a wide range of voice qualities. If the type of voice quality in the signal is contained within a smaller frequency range, high-frequency compensations may be less for that particular signal. For example, if it is known that only male voices will be contained in the signal, a considerable amount of high-frequency compensation may be eliminated.

Compensation for music is much more difficult since the frequency range of music varies so greatly. Frequency analysis varies from one musical selection to another and according to each individual's taste in frequency range reproduction. Also, it is necessary to consider the frequency limitations of the sound reproducing equipment since such large frequency spectrums are involved. Many small table radios have a high-frequency cutoff below 5,000 cycles per second while the more demanding listeners may have radios with a high-frequency cutoff above 20,000 cycles per second. In the first case, frequency compensation for noise frequency components above 5,000 cycles per second would cause overcompensation, while in the second case frequency compensation for noise frequency components up to 20,000 cycles per second may be needed. Both the type of reproduced music and the equipment reproducing the music will affect the frequency range of compensation needed.

Another important consideration is the purpose of the reproduced music. Most people will listen to a voice with the intention of understanding what is being said, but often music is played for mere background effect and maintaining intelligibility might not be critical or even desirable. Experimental results have shown it is possible to compensate for noise which masks music, but the remainder of this report will be concerned with voice masking because its effects are more critical and lend themselves more readily to study.

Once the filtering has been performed to allow proper weighting of noise frequencies, it remains to compare the level of the interfering noise with the level of the signal being listened to in order to determine if any increase in output volume is necessary in order to compensate for the noise. Detecting the noise level in the listening area may be accomplished by placing a microphone in the vicinity of the listener. The microphone, however, will pick up both the noise and the speaker signal output. If the speaker signal output is not subtracted from the microphone output, the speaker signal output would be mistaken for noise, thus creating a feedback situation where the louder the speaker output becomes, the more apparent noise would be present, causing the speaker output to be increased even more.

A block diagram showing the necessary circuitry to achieve compensation is shown in Figure 7. Two inputs are needed to determine the noise level, one input from the microphone which detects the signal plus the noise, and the other input from the main audio amplifier stage which contains only the signal. After filtering, d-c voltages proportional to the two inputs are subtracted from each other to give the noise level. The output voltage representing the noise level is used to control the automatic-gain-control stage of the main amplifier, altering the amplifier gain appropriately for the various noise levels. In order to have the output of the subtraction circuit indicate only the noise level, it is necessary to have, with no noise present, complete subtraction with no residue for all frequencies amplified in the main amplifier.

In order to have complete signal subtraction for all frequencies, it is necessary that the frequency response of the two channels be matched. It should be noted that the audio input channel is connected directly to the main amplifier, while the signal in the other channel travels through any final stages, output transformers, the speaker, the acoustical response of the listening area, into the microphone, and through the microphone amplifier and filter weight circuit. Since the signal in the microphone channel goes through many more stages

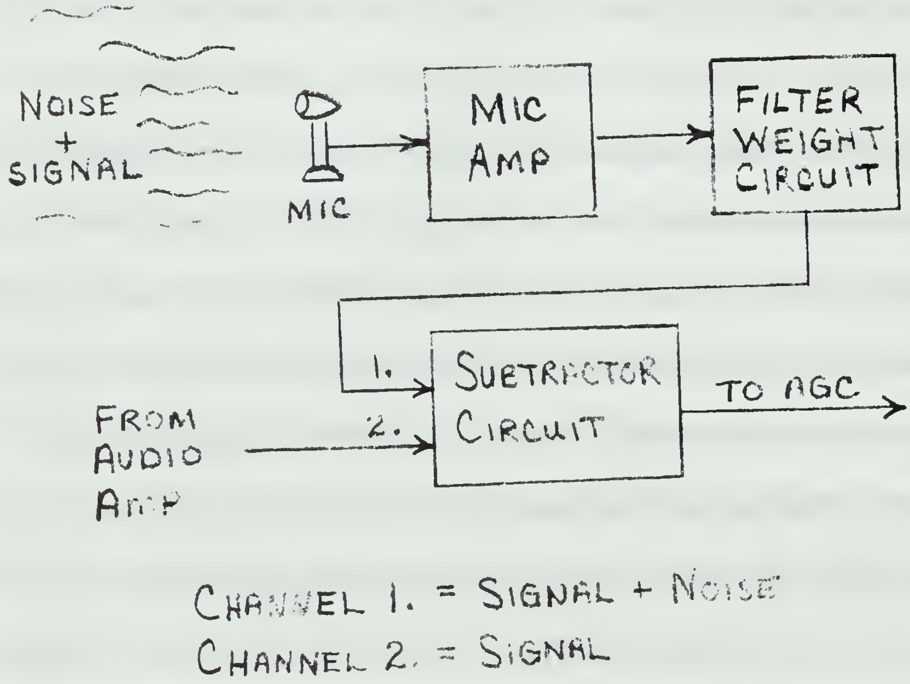


FIG. 7 - BLOCK DIAGRAM OF NOISE COMPENSATION CIRCUITRY

of processing, the frequency response of this channel will not be as broad as that of the audio input channel. In order to match the frequency response of the two channels then, it is necessary to limit the high-frequency and low-frequency responses of the audio channel to match that of the microphone channel. Since the major factor determining the frequency response in the microphone channel is the filter weight circuit, it is necessary to place an identical filter in the audio input channel. Experimental results show that with matching filters in each channel, not only will the noise frequencies be weighted as desired, but also the high-frequency and low-frequency responses of the two channels will be adequately matched.

Even with the high-frequency and low-frequency responses of the two channels matched, any unmatched "holes" or resonant peaks in the intermediate response of either channel will produce an output from the subtractor circuit, resulting in an erroneous AGC noise signal be being generated when no noise is in fact present. Such unmatched intermediate responses could be caused by holes or resonant peaks in the frequency response characteristics of the microphone or speaker and by reflection and refraction patterns in the acoustical response of the listening area. In order to minimize these channel differences, one should use good microphones and speakers with flat frequency response characteristics and should place the microphone in a good sampling area. Experimental results of tests in which these principles were applied will be discussed later.

Although theoretically it is possible to match even the intermediate-frequency response of the two channels in addition to the high-frequency and low-frequency responses, such circuitry would be extremely complicated and expensive to build, and once such circuitry were adjusted, merely moving objects in the listening area or changing the placement of the speaker or microphone would completely change the acoustical response of the environment and necessitate complete readjustment to achieve the close match. Luckily, perfectly matched channels are not necessary, and a certain degree of irregular response will not critically affect the operation of the circuit. There are two reasons for this. First, most signals heard will contain a smear of frequencies. The AGC circuitry will respond to the average power of all of the frequencies, so a few distorted individual frequency components will not be critical. Secondly, the system should be designed to slightly overcompensate so that small variances in the output level will not affect the intelligibility because of the residual gain above the threshold detection level. This, in effect, is accomplished anyway by the fact that most listeners will initially set the volume level considerably above that necessary for mere threshold detection. Considering the variances of the two channels, matching filter-weight circuits should adequately match the frequency responses of the two channels.

The outputs from the two filters cannot be directly subtracted because of the different amounts of phase shift induced in the two channels. To eliminate the phase-shift problem, each signal is rectified and then the two d-c signals are subtracted. The result of the subtraction is a d-c signal which is proportional to the noise level in the environment. This signal is used to control the gain of the AGC stage in the main amplifier in a manner which will maintain intelligibility of the original signal. Figure 8 shows the complete circuitry discussed so far.

In order to test the performance of this circuitry, a noise-level determination unit, the part of the circuit enclosed in the dotted line in Figure 8, was designed. This unit was then tested in a communications receiver utilizing the AGC circuitry of that receiver, and the results are reported in Chapter 3. The remaining part of this chapter will be concerned with the design of the noise-level determination unit, the final design of which is shown in Figure 9.

For test purposes, a dynamic microphone, Monarch TM-25, was selected because it was relatively inexpensive and the frequency response, 80 to 10,000 cycles per second, was extremely flat. Actually, for use with voice signals, the high-frequency range is not critical since the filter weight circuits will limit the response to frequencies below 5,000 or 6,000 cycles per second. The flat response is the critical characteristic. For music control, a greater frequency range is desirable as indicated previously. The TM-25 was found to be adequate for both music and voice; however, a much lower quality microphone would be satisfactory for voice control.

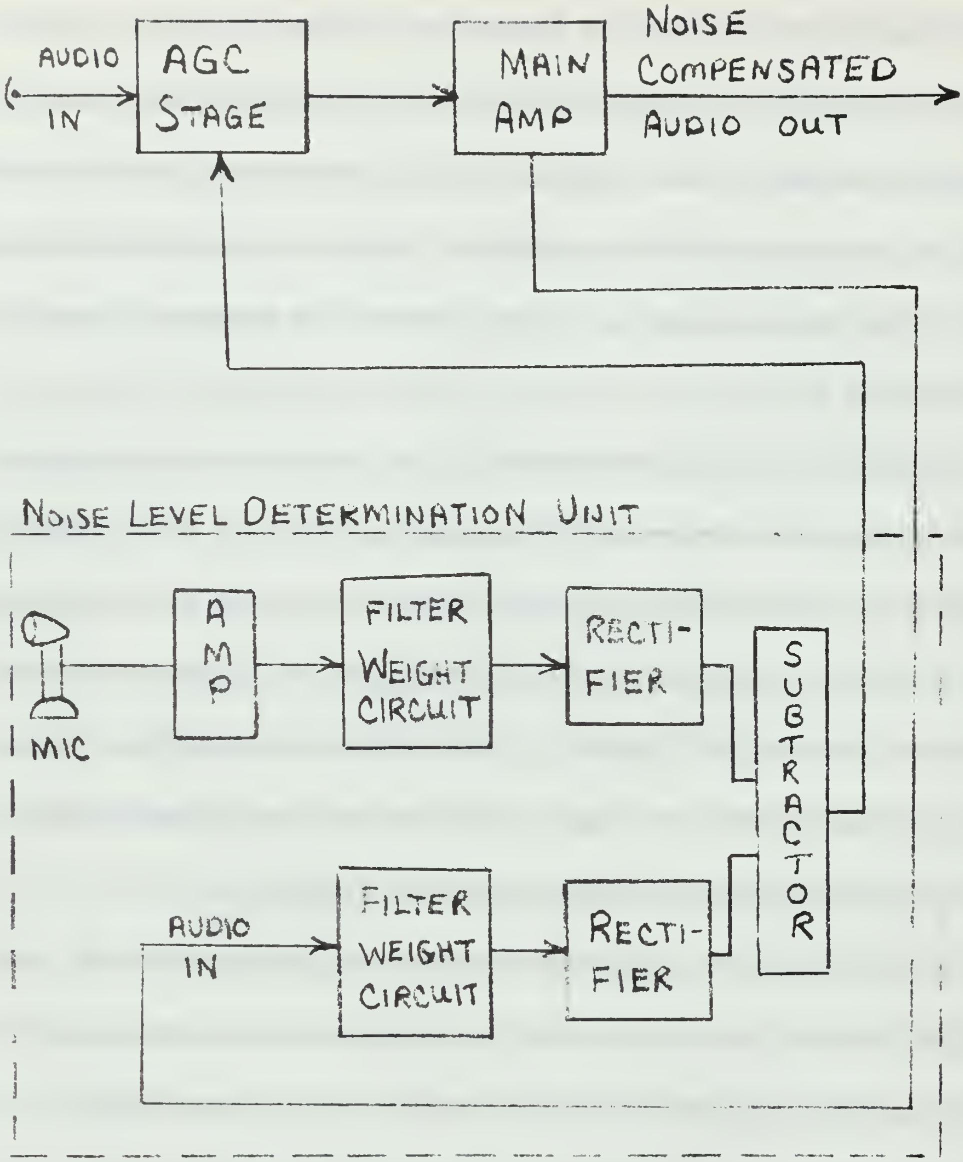
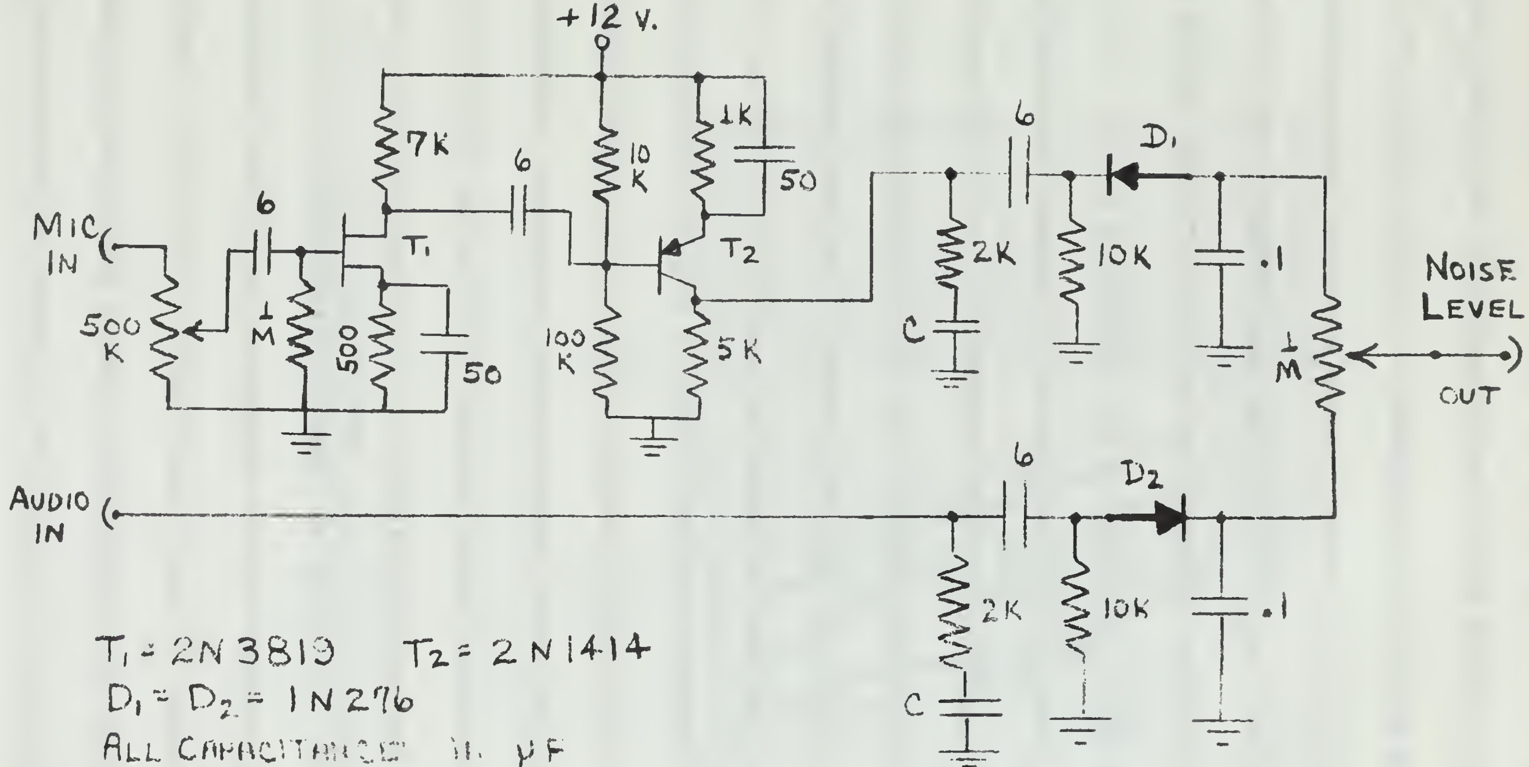


FIG 8.— BLOCK DIAGRAM OF NOISE COMPENSATION [illegible]



T_1 = 2N3819 T_2 = 2N1414
D_1 = D_2 = 1N276
All capacitances in μF
Resistances in Ω

Fig 9.- Noise Level Determination Unit

The audio amplifier in the microphone channel should provide enough gain for the desired noise-detection sensitivity and have a moderately high output impedance to match the filters and rectifiers. No amplification is needed in the audio input channel if the tap point is chosen late enough in the main amplifier circuitry to provide enough power. The power requirements will be discussed later in the chapter. The audio input impedance to this channel is high.

The matched filters were initially designed to match both the high-frequency and low-frequency responses of the two channels; however, experimental results showed that only the high-frequency responses of the two channels need be matched. The low-frequency responses of the two channels, although not necessarily matched, were similar enough to provide adequate control at the low-frequency portion of the spectrum where noise frequencies are seldom found anyway.

For matching the high-frequency responses, a capacitor and resistor in series placed in parallel across the input to the rectifier provides the necessary filtering. The placement of the capacitors, labeled "C", is shown in Figure 9. Figure 10 shows the experimental results of rectifier output voltage as a function of frequency for the various values of capacitance shown on the curves. Though the above simple filter only matches the cutoff or half-power frequencies and not necessarily the slope of the response, it was found to be more than adequate.

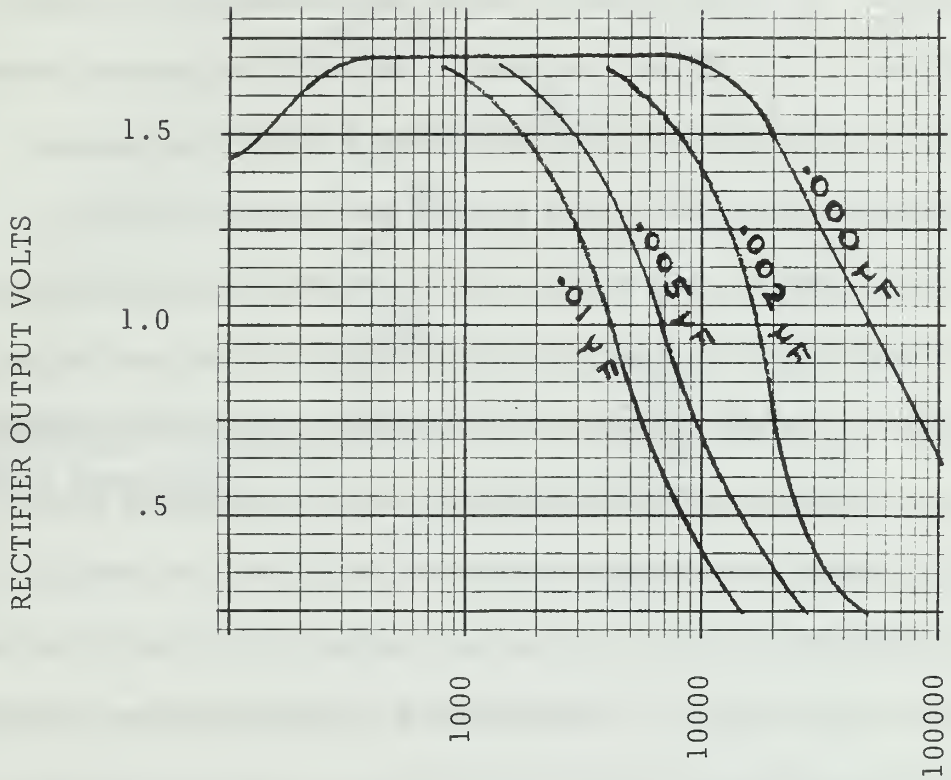


FIG. 10 - RECTIFIER OUTPUT VS FREQUENCY FOR VARIOUS
FILTER CAPACITORS

For voice control, a capacitance value of between .002 and .005 mf was found to be satisfactory, while for music control, the high-frequency response need be limited only by the necessity to match channels. The exact values of capacitance depend on the type of noise interference, but, for reasons previously explained, enough leeway exists in the use of the circuitry to allow a range of acceptable capacitance values which will work for most types of noise.

The rectifiers used between the filters and the subtraction circuit are simple diode rectifiers, shown in Figure 9. Note that the d-c voltages of the two channels are of opposite polarity. This allows the use of the balancing potentiometer as the subtracting circuit. With no noise present the potentiometer is adjusted by sending a signal through the main amplifier and varying the arm tap of the potentiometer until there is no output voltage. As long as there is no noise in the listening area, the ratio of the outputs of the two channels will remain constant for varying main amplifier signals and no output voltage will be generated. As soon as noise is detected by the microphone, the ratio of the two channel outputs will not be the same, thus the voltage at the tap arm will no longer be zero and an AGC voltage will be generated.

The action of the potentiometer subtractor may be shown mathematically. Let Vm and Va represent the voltage outputs from the microphone and amplifier channel respectively, and let Vo represent the output voltage from the potentiometer arm. Let R1 and R2 represent

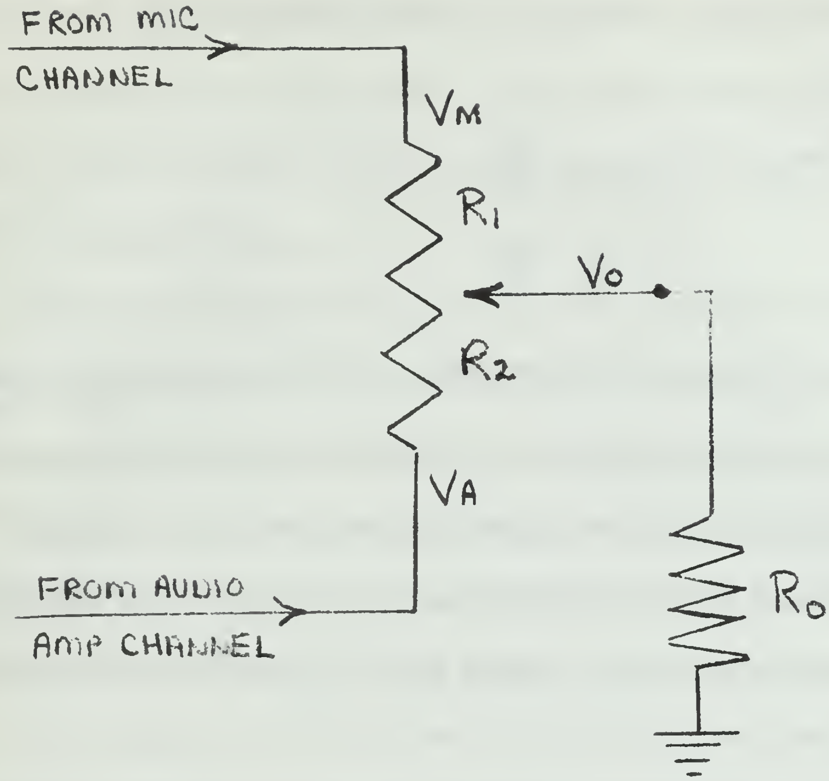


FIG. 11 - POTENTIOMETER SUBTRACTOR

the resistances of the potentiometer as shown in Figure 11, and let Ro represent the output resistance. Using Kirchoff's laws,

$$Vo = \frac{Vm + Va \left(\frac{R1}{R2}\right)}{1 + \frac{R1}{R2} + \frac{R1}{Ro}}$$

With Vm and Va of opposite polarity, Vo = 0 as long as $\frac{R1}{R2} = \frac{Vm}{Va}$, which is where the potentiometer is adjusted for operation with no noise present. As long as the above ratio is maintained and no noise is present, the output AGC voltage will be zero. However, as soon as the microphone detects noise, $|Vm| > |Va| \left(\frac{R1}{R2}\right)$, and an AGC voltage is generated.

In order to maximize Vo, which maximizes the sensitivity of the subtractor, Ro should be as large as possible to reduce the $\frac{R1}{Ro}$ term in the denominator. If this term is made negligible, then

$$\frac{dVo}{dVm} = \frac{1}{1 + \frac{R1}{R2}}$$

If R2 is much greater than R1, maximum sensitivity can be approached. But $Va = \frac{R2}{R1} Vm$ so for maximum sensitivity Va needs to be much larger than Vm, which itself must be large enough to give the desired voltage output. With vaximum sensitivity, dVo = DVm.

A differential amplifier could be used in place of the potentiometer subtractor, but this would require that the two signals be of equal amplitude with no noise present. The potentiometer subtractor not only avoids this requirement, but it is very easy to implement and adjust, and is inexpensive.

The signal-strength input to the audio channel should be as strong as possible to give maximum sensitivity. Generally, a voltage sample from a point in the final amplifier stage before the speaker output is adequate. The amount of amplification needed in the microphone channel depends on the output level of the microphone and the desired strength of the noise-indication voltage at the output of the subtractor.

Two important time constants in the circuit are the "attack" and "release" time constants. The attack time constant is the delay time constant between the time the noise is generated and the time the gain of the signal amplifier is increased. This time constant is mainly determined by the value of the filter capacitors at the input to the potentiometer subtraction circuit. The release time constant is the delay time constant between the disappearance of the noise and the quieting of the signal amplifier. This time constant is mainly determined by the total value of the resistance across the extreme ends of the potentiometer. The value of this resistance will determine how quickly the two filter capacitors, of opposite polarity, will discharge.

The release time constant is also dependent on the input impedance of the circuit to which the arm of the potentiometer is attached. If this impedance is high enough, as it should be for reasons previously explained, the release time constant is almost totally dependent on the overall potentiometer resistance. If this impedance is fairly low, the release time constant of each channel will depend on the combined effect of the value of the arm impedance and the position at which the arm rests. Under these circumstances, each channel feeding the potentiometer will have different release time constants. This unbalancing effect is eliminated when the arm impedance is high enough.

The output of the noise-level determination circuit shown in Figure 9 is a d-c voltage proportional to the noise level in the listening area which would interfer with the speaker output. The output signal from this circuit, further processed if necessary, is the basic signal used to control the AGC of the amplifier in a manner which compensates for the noisy environment.

IV. EXPERIMENTAL RESULTS

In order to test the principles used in designing the circuitry for automatic volume compensation in noisy enviroments and to test the performance of the noise determination unit, the unit was tested in conjunction with a commercial communications receiver and the system's performance was observed under various environmental conditions. The communications receiver used was a Hallicrafter Model SX-130. The receiver was modified so that the AGC designed into the receiver was controlled by the external noise determination unit instead of by the internal control voltages.

Figure 12 is a reproduction of a large portion of the schematic diagram of the SX-130, [12]. V5 generates the internal AGC voltage which controls the gain of V1, V3A and V4. The V5 plate side of R31 was disconnected, and the AGC voltage from the noise determination unit was connected to the internal AGC circuit through R31. The audio output signal from the receiver, to be used in the audio input channel of the noise level determination unit, was taken directly from the plate of V7, pin 6, shown in Figure 13, [13]. These were the connections made between the receiver and the noise-level determination unit discussed in Chapter 2.

By disconnecting R31 from the internal AGC control voltage, the normal stability of carrier strength in the receiver is undermined. However, for test purposes the RF Gain Control can be manually set to the desired "no noise" value for signal carriers of various strengths.

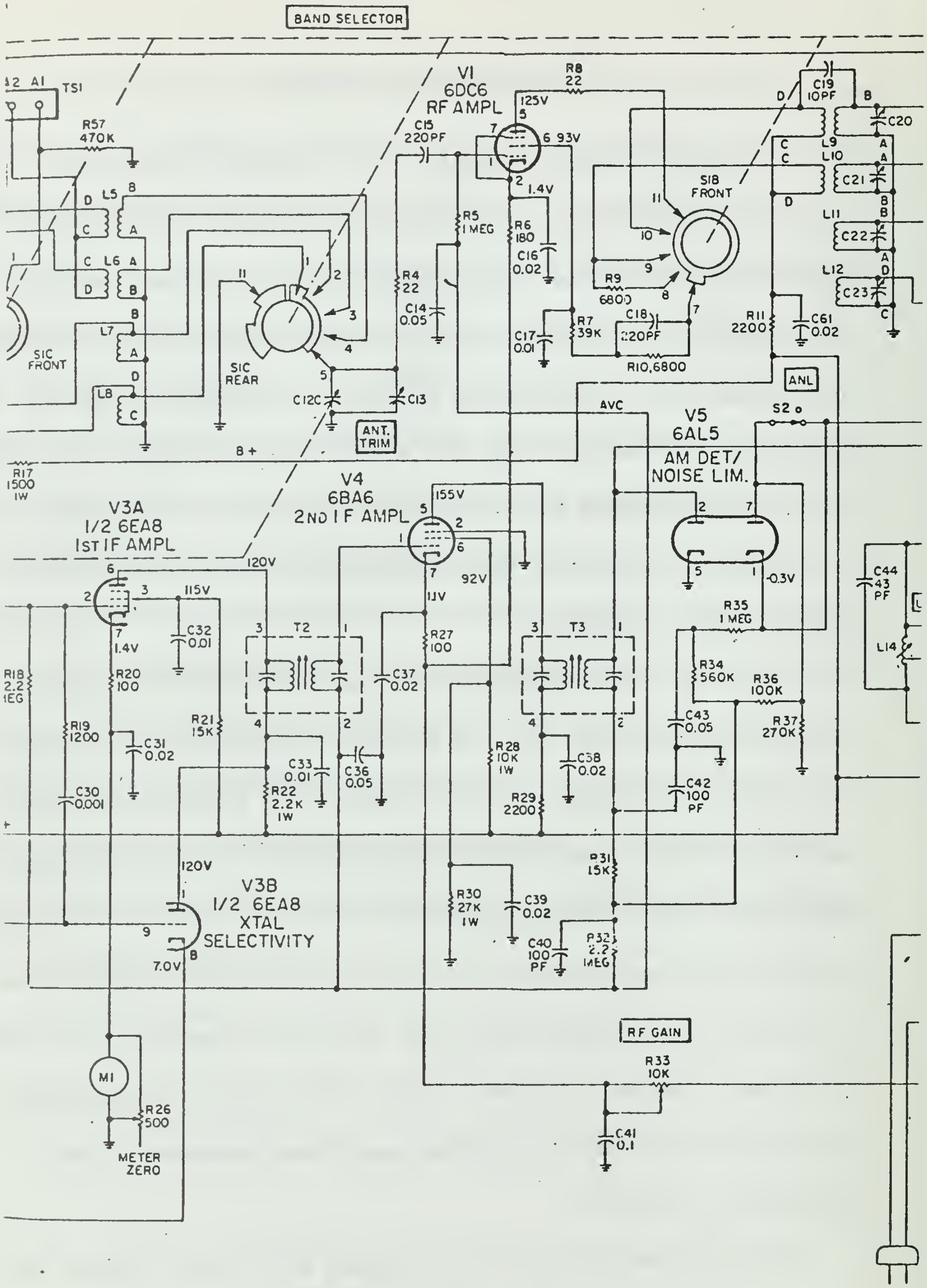


FIG. 12 - PARTIAL SCHEMATIC OF SX-130 FROM HALLICRAFTER'S INSTRUCTION MANUAL

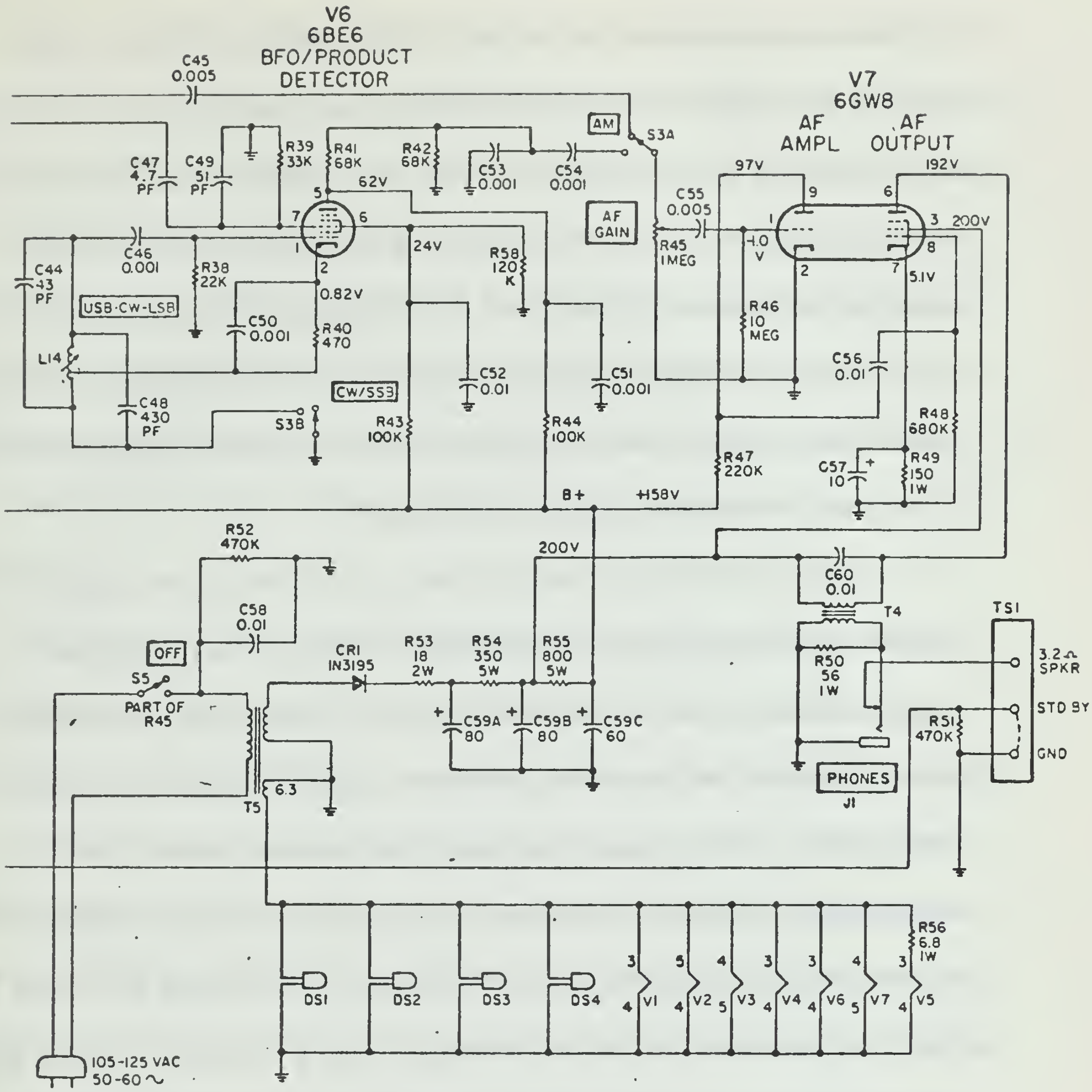


FIG. 13 - PARTIAL SCHEMATION OF SX-130 FROM HALLICRAFTER'S INSTRUCTION MANUAL

In a receiver which has an AGC circuit designed with a delay voltage, the output from the noise-level determination unit could be used to control the delay bias. Using this method of incorporating the external AGC control, the internal AGC control of the receiver could be left functioning and both the original and the added AGC controld could function simultaneously. The SX-130 receiver did not have a delayed AGC circuit and therefore, for experimental purposes, the internal AGC control was disconnected.

In order to match the dynamic ranges of the AGC unit output voltages and the receiver internal AGC circuitry, the normalized receiver output power as a function of AGC voltage was determined. Figure 14 shows the relationship for the moderate range of receiver output power. The ordinate indicates the receiver output power normalized to a 10-volt peak sine-wave measurement at the plate of the final audioamplifier tube. The abscissa indicates the value of AGC voltage both on an absolute scale and a db scale. Actual AGC voltages are negative voltages.

From Figure 14 it can be seen that the slope of the response curve is negative. In other words, the more negative the control voltage, the lower the receiver audio output power. For this reason, the noise-generated AGC voltage must be positive. The receiver AGC circuitry is biased to operate at a particular point on this curve, and any noise-generated AGC voltage will be positive. When added to the AGC bias voltage this noise component will make the total AGC voltage less negative, causing the output power of the receiver to be increased.

It is seen from Figure 14 that quite a range of slopes are available to bias the operating point on. The no-noise bias point may be anywhere along this curve as long as the audio amplifier can be manually adjusted to give adequate output with no noise present, and as long as enough residual gain is left in the receiver to provide adequate AGC control. The first condition places a limitation on how low on the operating curve the receiver may be biased while the second condition places a limitation on how high on the operating curve the receiver may be biased.

For example, reference to Figure 14 will show that if the receiver is biased at -3 AGC volts the receiver output power may be increased only 3 db for the worst noise conditions, while if the receiver is biased at -5 AGC volts the output power may be increased over 10 db for the worst noise conditions. The latter condition, however, will not allow the same maximum amount of audio output power with no noise present because of the large degree of quieting in the AGC stage of the receiver.

Also to be considered in the selection of a biased operating point on the AGC curve is the desired slope of the curve over the AGC operating range. The slope of the curve will determine the incremental amount of power output gain for incremental increases in noise. Since the incremental changes of AGC voltage are a function of the gain in the microphone channel of the noise determination unit, the potentiometer in the microphone input channel will actually determine the amount of AGC voltage generated for various noise conditions. With this in mind,

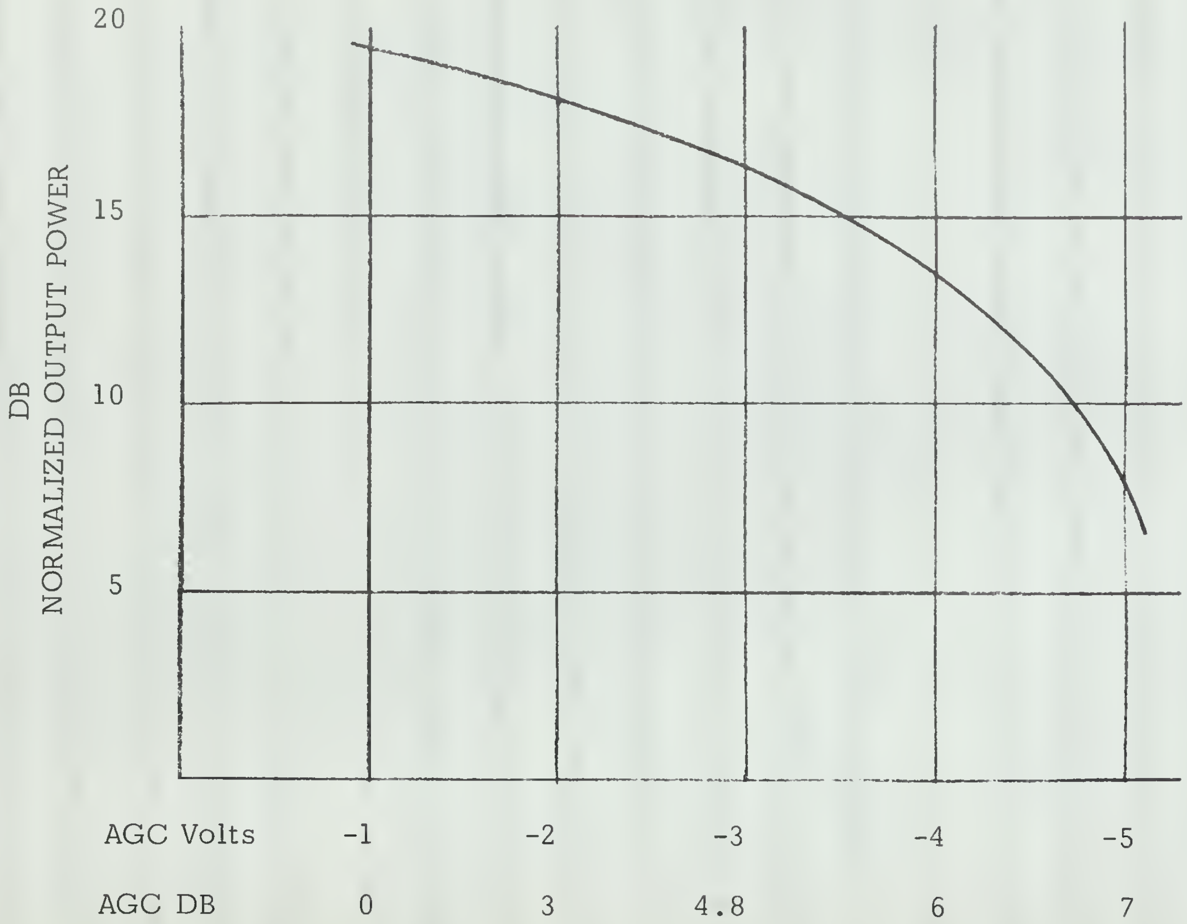


FIG. 14 - RECEIVER AGC CONTROL FUNCTION

any operating point on the curve may be chosen, and the incremental change of output power per incremental change of noise will be a function of the microphone potentiometer setting. It is more important, therefore, to choose the operating point according to the SHAPE of the curve rather than according to slope values. The numerical value of the slope may be changed by simply changing the microphone potentiometer setting.

The desired shape of the operating curve in the operating range may be determined from Figure 2. This figure shows the shift in the threshold hearing level for various intensities of masking tones. The frequency of the masking tone is shown near the top of each graph and the frequency of each tone being masked is shown on the individual curves of each graph. The shape of the curve in Figure 14 over the operating range of AGC voltage should be similar to the shape of the threshold shift curve of Figure 2 for the worst possible case of masking. If these frequencies representing the worst possible cases of masking are compensated for, overcompensation will result for any other frequencies exhibiting a lesser vulnerability to masking.

The shape of the masking curve for the worst case of masking shown in Figure 2 is a straight line, the straight line labeled 850 cycles per second in the left curve and the line labeled 2350 cycles per second in the right curve. The masking curves for the other frequencies on each graph deviate from a straight line; but if the straight-line case is compensated for, the other cases of masking will be more than compensated for, since the straight-line case represents the maximum degree of masking effectiveness. The desired shape of the curve of Figure 14 over the AGC operating range would then be a straight line.

Since Figure 14 does not offer a straight-line operating curve over any extended range, two cases of straight-line approximation were evaluated. Figure 15 is a reproduction of Figure 14 with the two cases of straight-line approximation superimposed. An initial bias point of -41/2 volts was chosen. Since the setting of the potentiometer in the microphone channel determines the value of AGC control voltage generated for a given amount of noise, the values of potentiometer settings to approximate both case 1 and case 2 were determined and the performance of the control circuit for each setting was analyzed.

With the microphone potentiometer set for the idealized case 1, the response of the output power in db to changes of AGC voltage is fairly linear from about 11 db to about 14 db of output power. This linearity results in the desired straight-line response discussed with reference to Figure 15 and Figure 2. Although the actual portion of the total curve of Figure 15 which can be considered linear for this bias point is a small percentage of the total curve over which the AGC voltage might operate, it does represent a doubling of output power, or an increase of 3 db, over a fairly linear portion of the curve in Figure 15. If noise conditions demand an operating point much higher on the idealized curve of case 1, the output of the receiver will not be

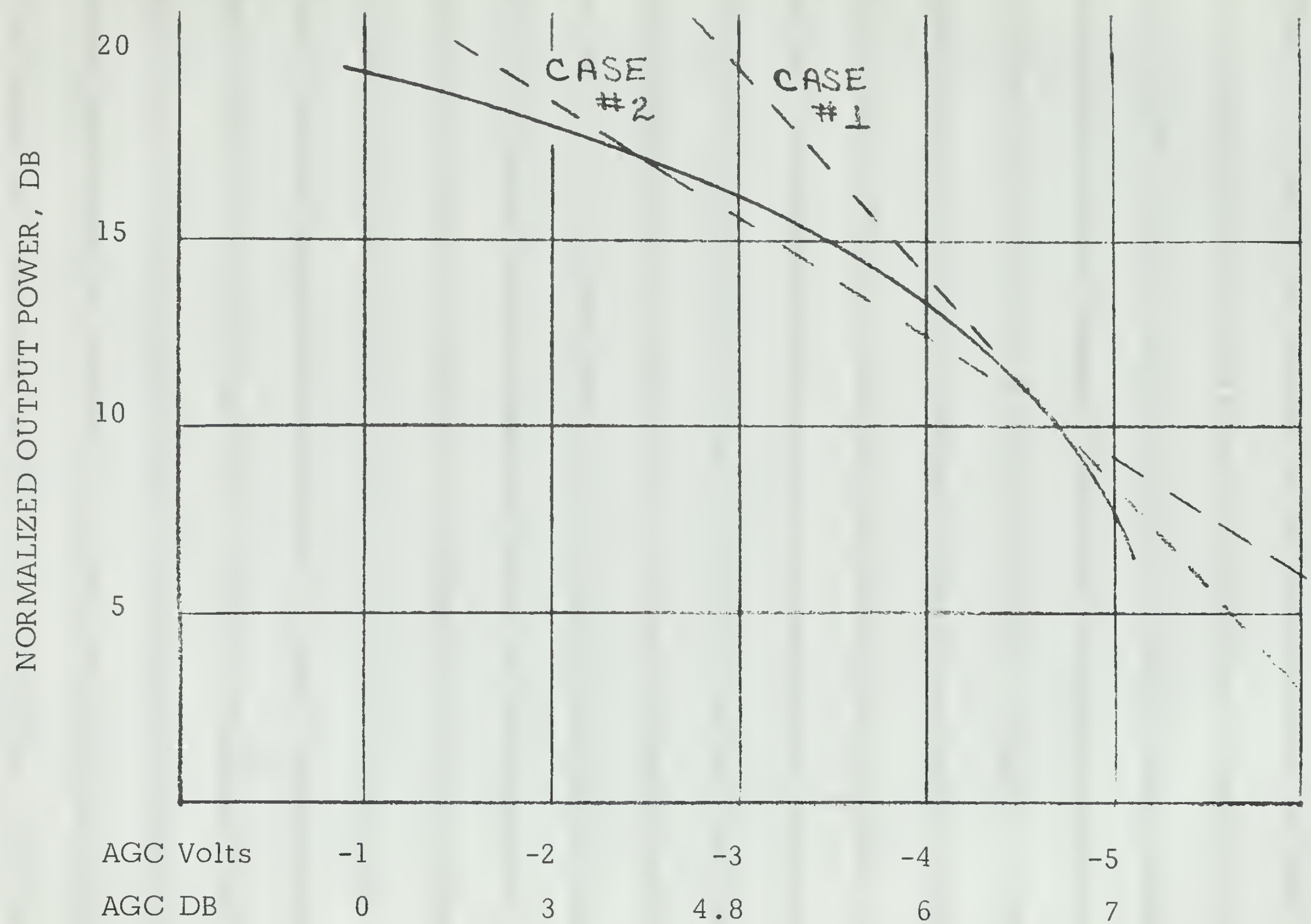


FIG. 15 - AGC FUNCTION WITH SUPERIMPOSED IDEAL OPERATING CURVES

adequate to compensate for the noise because of the deviation of the actual curve from the idealized curve. For output power increases of up to 3 db, however, this particular adjustment of the microphone potentiometer works very well.

With the microphone potentiometer set for the desired idealized case 2, the actual system-response deviation was much greater. From a receiver output power of 11 db to a power output of 17 db, the actual output power was greater than that demanded by the noise conditions, as Figure 15 shows. Too much output power does not degrade intelligibility of the output signal, however. It is not until the output power falls below the required power that intelligibility is degraded. Potentiometer setting for case 2 maintains intelligibility in spite of noise for a receiver power output range of from 11 db to 17 db, while a potentiometer setting for case 1 maintains signal intelligibility for receiver power outputs from 11 db to only 14 db.

In order to set the microphone potentiometer correctly, it should be adjusted for signal intelligibility under maximum noise conditions. This setting will determine the upper point of intersection for the idealized curve of case 2 and the actual operating curve of Figure 15. For smaller amounts of noise, more than enough compensation will be made. Since the output power level will probably not be at a minimum threshold hearing level anyway, but will be at a comfortable power level above this, intelligibility will still be maintained for noise levels slightly above the noise level setting.

The noise-level determination unit, as shown in Figure 9, must be modified before it can provide the AGC voltage for the internal AGC circuitry of the SX-130. The SX-130 receiver requires a steady-state negative d-c bias voltage with a superimposed positive voltage representing the noise level. The positive component makes the negative bias less negative causing the receiver audio output power to increase.

The negative bias for the receiver AGC circuit may be obtained by adding another transistor stage to the noise-level determination circuitry shown in Figure 9. This addition is shown in Figure 16. The negative d-c bias point, which determines the no-noise bias point of Figure 15, may be varied by changing the value of the base bias resistor of T3. Although the d-c signal representing the noise level is negative at the output from the potentiometer, T3 inverts this signal and gives the correct combination of bias voltage and noise-level voltage and in addition provides additional gain in the AGC circuitry.

For the component values shown in Figure 16, the bias point shown in Figure 15 is obtained. With the microphone potentiometer set to approximate case 2 and using the noise source #2 of Table 1, the response shown in Figure 17 was obtained. Figure 17 is a comparison of the speaker output for various levels of noise. The large sine wave in each picture is the audio output signal on the plate of the final audio amplifier tube in the receiver. The remaining frequency components

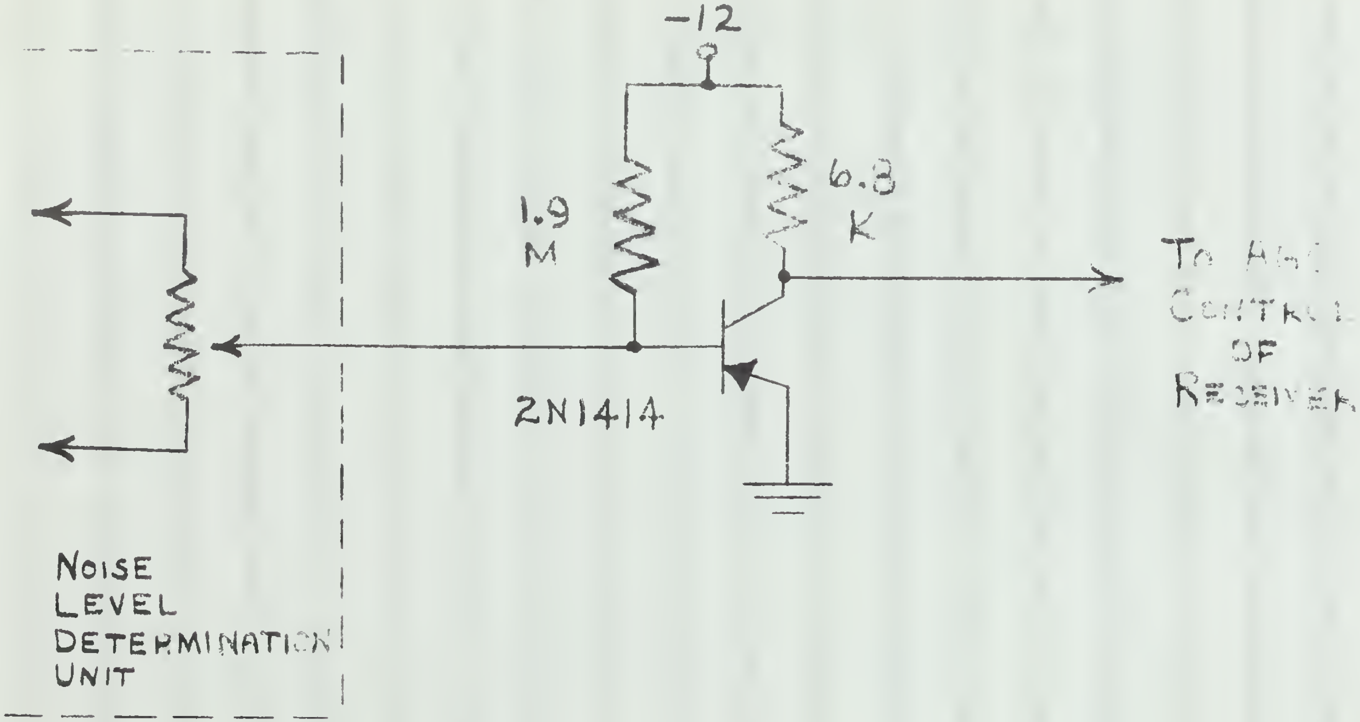


FIG. 16 - STAGE ADDED TO UNIT FOR OPERATION WITH SX-130

are the noise as it appears at the collector of T2 in Figure 9. The four pictures show the audio output signal superimposed on the noise signal for four different levels of noise.

If the microphone used to sense the environmental noise level is not omnidirectional, the placement of the microphone will determine the directional pattern of sensitivity of the system. In this case, the microphone should be placed as close as possible to the listener and the listening environment so that an accurate sample of what the listener actually hears is obtained. If the microphone is too far from the listener the system will compensate for noises in the immediate vicinity of the microphone which may not be the necessary compensation for the noises in the vicinity of the listener.

There are two problems to consider when determining how to place a directional microphone. The first problem is to place the microphone in such a way as to minimize the coupling between the output speaker and the microphone which could cause feedback. This can be accomplished by placing the microphone between the listener and the speaker and aiming the microphone in the direction of the listener. If the microphone is placed in this manner, the speaker output will be in a direction of minimum microphone gain and the microphone gain may be greatly increased to give greater sensitivity without danger of feedback. Because of this decoupling advantage which accrues from the correct placement of a directional microphone, the use of a directional microphone is definitely preferred to the use of an omnidirectional microphone.

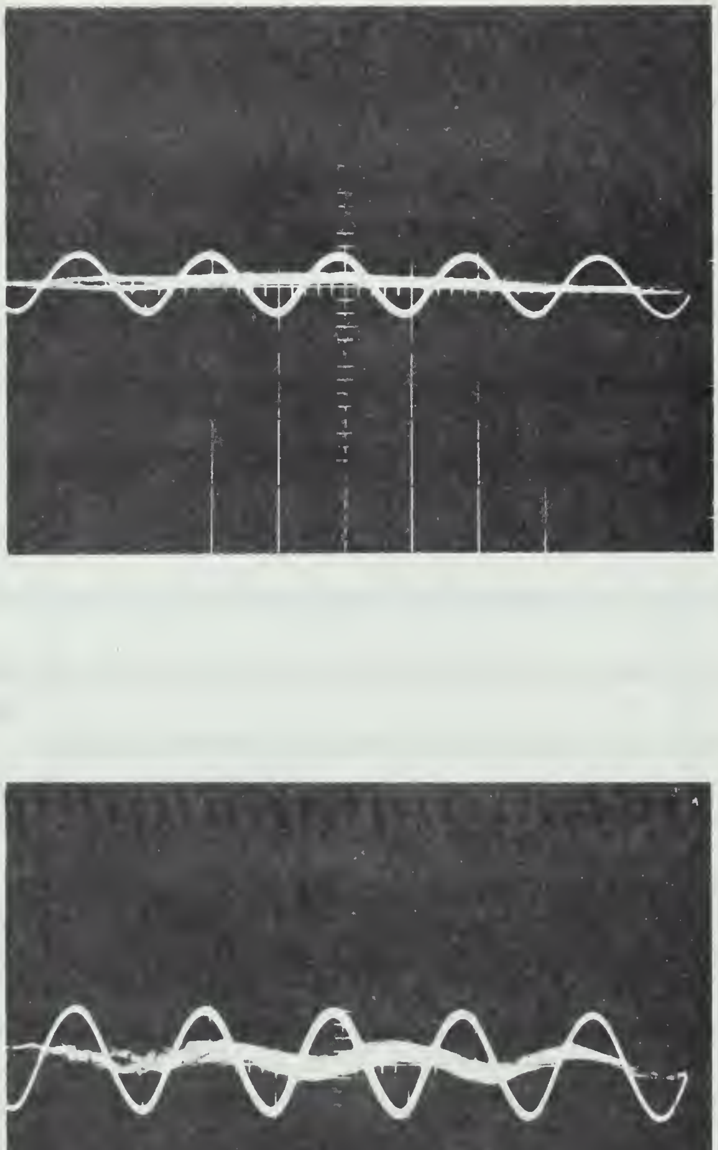


FIG. 17A - RECEIVER OUTPUT VS ENVIRONMENTAL NOISE

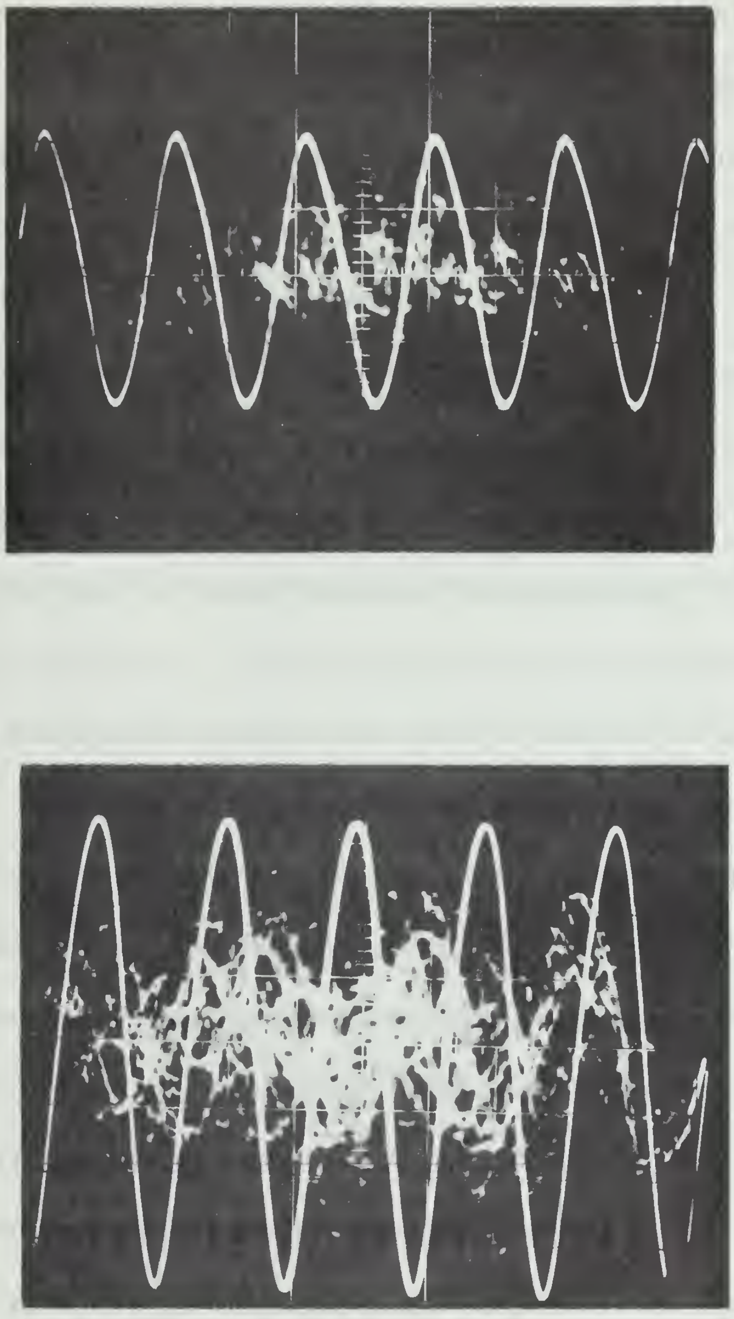


FIG. 17B - RECEIVER OUTPUT VS ENVIRONMENTAL NOISE

The second problem arises from using a directional microphone and placing it in the manner described above. With the microphone placed between the listener and the speaker and aimed at the listener, it is most sensitive to noises on the opposite side of the listener from the speaker and it is least sensitive to noises on the same side of the listener as the speaker. This results in the system being more sensitive to stereo masking, where the signal and the noise come from different directions to the listener, than to monaural masking, where the signal and the noise come from the same direction to the listener. This effect is exactly opposite from the desired effect as described in Chapter 1. The microphone should be more sensitive to monaural masking since this type of masking is much more effective than stereo masking for any given signal and noise level.

Experimental results have shown that when the system is operated in a confined area where the speaker and the microphone are tightly coupled acoustically, such as in an automobile, it is more important to place the microphone for a minimum amount of feedback coupling and to sacrifice the desired stereo-monaural effect. In larger areas where there is greater acoustical separation between the speaker and the microphone, such as in a large office or work area, a maximum amount of decoupling is not necessary and some degree of the desired stereo-monaural effect can be achieved by having the microphone aimed slightly towards the speaker. It should be remembered that, as mentioned before, the volume of the speaker output will be greater than that just necessary for threshold detection, so inaccuracies in stereo-monaural noise weighting will not be critical

V. CONCLUSIONS

Experimental results have shown that the technique of Automatic Volume Control for Noisy Environments as described in this thesis is a practical method for maintaining the intelligibility of speaker output signals in spite of interfering environmental noises. When used with existing automatic control circuitry this technique could provide the final compensation necessary to guarantee signal readability.

LIST OF REFERENCES

1. Harold M. Kaplan, Anatomy and Physiology of Speech (New York, 1960), pp. 302-306.

2. Harold M. Kaplan, Anatomy and Physiology of Speech (New York, 1960), pp. 207-267.

3. Harvey Fletcher, Speech and Hearing in Communications (New Jersey, 1953), Chapter 10.

4. Harvey Fletcher, Speech and Hearing in Communications (New Jersey, 1953), pp. 153-155.

5. Harvey Fletcher, Speech and Hearing in Communications (New Jersey, 1953), pp. 156-157.

6. Harvey Fletcher, Speech and Hearing in Communications (New Jersey, 1953), p. 157.

7. Harvey Fletcher, Speech and Hearing in Communications (New Jersey, 1953), p. 165.

8. Harvey Fletcher, Speech and Hearing in Communications (New Jersey, 1953), pp. 200-206.

9. Harvey Fletcher, Speech and Hearing in Communications (New Jersey, 1953), pp. 157-159.

10. Harvey Fletcher, Speech and Hearing in Communications (New Jersey, 1953), pp. 158-159.

11. American Radio Relay League, The Radio Amateur's Handbook (Connecticut, 1966), p. 247.

12. Hallicrafter, Operation Manual For SX-130, supplied with receiver.

13. Hallicrafter, Operation Manual For SX-130, supplied with receiver.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	20
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	NAVELEX Systems Command Department of the Navy Washington, D. C. 20390	1
4.	Professor D. B. Hoisington Department of Electronics and Communications Naval Postgraduate School Monterey, California 93940	1
5.	LTjg Charles R. McGough, USN Apartment E-2 1760 Clairmont Road Decatur, Georgia 30033	1

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Naval Postgraduate School Monterey, California 93940	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

Automatic Volume Compensation For Noisy Environments

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Master's Thesis; June 1969

5. AUTHOR(S) *(First name, middle initial, last name)*

Charles Randall McGough

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
June 1969	63	13

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO.	
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	

10. DISTRIBUTION STATEMENT

Distribution of the document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Naval Postgraduate School Monterey, California 93940

13. ABSTRACT

The importance of environmental noises which interfere with the audio output from the loudspeaker of an intercom or radio communications receiver has led to the study of sound masking and noise interference. The results are applied in the development of an automatic control system which compensates for environmental noises by maintaining the speaker output at a prescribed listening level above such interference. Such a system could be used to maintain the readability of communications in spite of a noisy environment such as on the bridge or in the engine room of a ship.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6811

Security Classification

A-31408

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Environmental noise Automatic compensation						